



# Potential sources of variability in mesocosm experiments on the response of phytoplankton to ocean acidification

Maria Moreno de Castro<sup>1</sup>, Markus Schartau<sup>2</sup>, and Kai Wirtz<sup>1</sup>

<sup>1</sup>Helmholtz-Zentrum Geesthacht, Centre for Materials and Coastal Research, Geesthacht, Germany

<sup>2</sup>GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

*Correspondence to:* Maria Moreno de Castro (maria.morenodecastro@outlook.com)

Received: 9 March 2016 – Discussion started: 8 April 2016

Revised: 6 March 2017 – Accepted: 13 March 2017 – Published: 6 April 2017

**Abstract.** Mesocosm experiments on phytoplankton dynamics under high CO<sub>2</sub> concentrations mimic the response of marine primary producers to future ocean acidification. However, potential acidification effects can be hindered by the high standard deviation typically found in the replicates of the same CO<sub>2</sub> treatment level. In experiments with multiple unresolved factors and a sub-optimal number of replicates, post-processing statistical inference tools might fail to detect an effect that is present. We propose that in such cases, data-based model analyses might be suitable tools to unearth potential responses to the treatment and identify the uncertainties that could produce the observed variability. As test cases, we used data from two independent mesocosm experiments. Both experiments showed high standard deviations and, according to statistical inference tools, biomass appeared insensitive to changing CO<sub>2</sub> conditions. Conversely, our simulations showed earlier and more intense phytoplankton blooms in modeled replicates at high CO<sub>2</sub> concentrations and suggested that uncertainties in average cell size, phytoplankton biomass losses, and initial nutrient concentration potentially outweigh acidification effects by triggering strong variability during the bloom phase. We also estimated the thresholds below which uncertainties do not escalate to high variability. This information might help in designing future mesocosm experiments and interpreting controversial results on the effect of acidification or other pressures on ecosystem functions.

## 1 Introduction

Oceans are a sink for about 30 % of the excess atmospheric CO<sub>2</sub> generated by human activities (Sabine et al., 2004). Increasing carbon dioxide concentration in aquatic environments alters the balance of chemical reactions and thereby produces acidity, which is known as ocean acidification (OA) (Caldeira and Wickett, 2003). Interestingly, the sensitivity of photoautotrophic production of particulate organic carbon (POC) to OA is less pronounced than previously thought. Several studies on CO<sub>2</sub> enrichment revealed an overall increase in POC (e.g., Schluter et al., 2014; Eggers et al., 2014; Zondervan et al., 2001; Riebesell et al., 2000), but other studies did not detect CO<sub>2</sub> effects on POC concentration (e.g., Jones et al., 2014; Engel et al., 2014) or primary production (Nagelkerken and Connell, 2015). General compilation studies that document controversial results are, e.g., Riebesell and Tortell (2011) and Gao et al. (2012).

In some experiments, the different treatment levels, i.e., different CO<sub>2</sub> concentrations, have been applied in parallel repetitions, also known as replicates or sample units. This was the case in several CO<sub>2</sub> perturbation experiments with mesocosms (Riebesell et al., 2008). Often, high variances are found in measurements among replicates of similar CO<sub>2</sub> levels (Paul et al., 2015; Schulz et al., 2008; Engel et al., 2008; Kim et al., 2006; Engel et al., 2005). It is this variance in data that reflects system variability, thereby introducing a severe reduction in the ratio between a true acidification response signal and the variability in observations. Ultimately, the experimental data exhibit a low signal-to-noise ratio.

Mesocosms typically enclose natural plankton communities, which is a more realistic experimental setup compared to

batch or chemostat experiments with monocultures (Riebesell et al., 2008). Along with this, mesocosms allow for a larger number of possible planktonic interactions that provide opportunities for the spread of uncontrolled heterogeneity. Moreover, physiological states vary for different phytoplankton cells and environmental conditions. For this reason, independent experimental studies at similar but not identical conditions might yield divergent results. The variability in data of mesocosm experiments is thus generated by variations of ecological details, i.e., small differences among replicates of the same sample, such as in species abundance, nutrient concentration, and metabolic states of the algae at the initial setup of the experiments. Differences of these factors often remain unresolved and might therefore be treated as uncertainties in a probabilistic approach.

To account for all possible factors that determine all differences in plankton dynamics is practically infeasible, which also impedes a retrospective statistical analysis of the experimental data. However, since unresolved ecological details might propagate over the course of the experiment, it is meaningful to consider a dynamical model approach to upgrade the data analysis. From a modeling perspective, some important unresolved factors translate into (i) uncertainties in specifying initial conditions (of the state variables), and (ii) uncertainties in identifying model parameter values. Here, we apply a dynamical model to estimate the effects of eco-physiological uncertainties on the variability in POC concentration of two mesocosm experiments. Our model describes plankton growth in conjunction with a dependency between CO<sub>2</sub> utilization and mean logarithmic cell size (Wirtz, 2011). The structure of our model is kept simple, thereby reducing the possibility of overparameterizing the mesocosms dynamics. The model is applied to examine how uncertainties in individual factors, namely initial conditions and parameters, can produce the standard deviation of the distribution of observed replicate data. Our main working hypotheses on the origins of variability in mesocosm experiments are the following:

- Differences among replicates of the same sample can be interpreted as unresolved random variations (named uncertainties hereafter).
- Uncertainties can amplify during the experiment and generate considerable variability in the response to a given treatment level.
- Which uncertainties are more relevant can be estimated by the decomposition of the variability in the experimental data.

For our data-supported model analysis of variability decomposition we consider the propagation of distributions (JCGM, 2008b) to seek potential treatment responses that are masked by the variability in observations of two independent OA mesocosm experiments, namely, the Pelagic Enrichment CO<sub>2</sub> Experiment (PeECE II and III). The central idea

is to produce ensembles of model simulations, starting from a range of values for selected factors. The range of values for these selected factors is determined so as the variability in model outputs does not exceed variability in observations over the course of the experiment. The margins of the variational range of each factor were thus confined by the ability of the dynamical model to reproduce the magnitude of the variability observed in POC. These confidence intervals describe the tolerance thresholds below which uncertainties do not escalate to high variability in the modeled replicates, and can serve as an estimator of the tolerance of experimental replicates to such uncertainties. This information can be important to ensure reproducibility, allowing for a comparison between the results of different independent experiments and increasing confidence regarding the effects of OA on phytoplankton (Broadgate et al., 2013).

## 2 Method

Potential sources of variability are estimated following a procedure already applied in system dynamics, experimental physics, and engineering (JCGM, 2008b). The basic principles of uncertainty propagation are summarized here using a six-step method (see Fig. 1). Steps 1 and 2 are described in Sect. 2.1 and comprise a classical model calibration (using experimental data of biomass and nutrients) to obtain the reference run representing the mean dynamics of each treatment level. In this way we found the reference value for the model factors, i.e., parameters and initial conditions. Steps 3 and 4, described in Sect. 2.2, include the tracked propagation of uncertainties by systematically creating model trajectories for POC, each one with a slightly different value of a model factor. In steps 5 and 6, we estimated the thresholds of the model-generated variability and the effect of the uncertainty propagation (also explained in Sect. 2.2).

### 2.1 Model setup, data integration, and description of the reference run

In this section, we describe the biological state that was used as reference dynamics. Our model resolves a minimal set of state variables insofar monitored during experiments that are assumed to be key agents of the biological dynamics. Model equations are shown in Table 1. Reference values of the parameters are shown in Table 2. An exhaustive model documentation is given in Appendix A. The model simulates experimental data from the Pelagic Enrichment CO<sub>2</sub> Experiment (PeECE), a set of nine outdoor mesocosms placed in coastal waters close to Bergen (Norway) during the spring seasons of 2003 (PeECE II) and 2005 (PeECE III). In both the experiments, blooms of the natural phytoplankton community were induced and treated in three replicates for the future, present, and past CO<sub>2</sub> conditions (Engel et al., 2008; Schulz et al., 2008; Riebesell et al., 2007, 2008). Experiment-

**Table 1.** States variables and their dynamics.

State variable	Dynamical equation	Ini. cond.	Units
Phytoplankton carbon	$\frac{d\text{Phy}_C}{dt} = (P - R - L) \cdot \text{Phy}_C$	2.5	$\mu\text{mol-CL}^{-1}$
Phytoplankton nitrogen	$\frac{d\text{Phy}_N}{dt} = V \cdot \text{Phy}_C - L \cdot \text{Phy}_N$	0.4	$\mu\text{mol-NL}^{-1}$
Nutrient concentration	$\frac{d\text{DIN}}{dt} = r \cdot \text{DH}_N - V \cdot \text{Phy}_C$	$8 \pm 0.5^*$ $14 \pm 2^{**}$	$\mu\text{mol-NL}^{-1}$ $\mu\text{mol-NL}^{-1}$
Detritus and heterotrophs C	$\frac{d\text{DH}_C}{dt} = L \cdot \text{Phy}_C - (s \cdot \text{DH}_C + r) \cdot \text{DH}_C$	0.1	$\mu\text{mol-CL}^{-1}$
Detritus and heterotrophs N	$\frac{d\text{DH}_N}{dt} = L \cdot \text{Phy}_N - (s \cdot \text{DH}_N + r) \cdot \text{DH}_N$	0.01	$\mu\text{mol-NL}^{-1}$

\* PeECE II, \*\* PeECE III

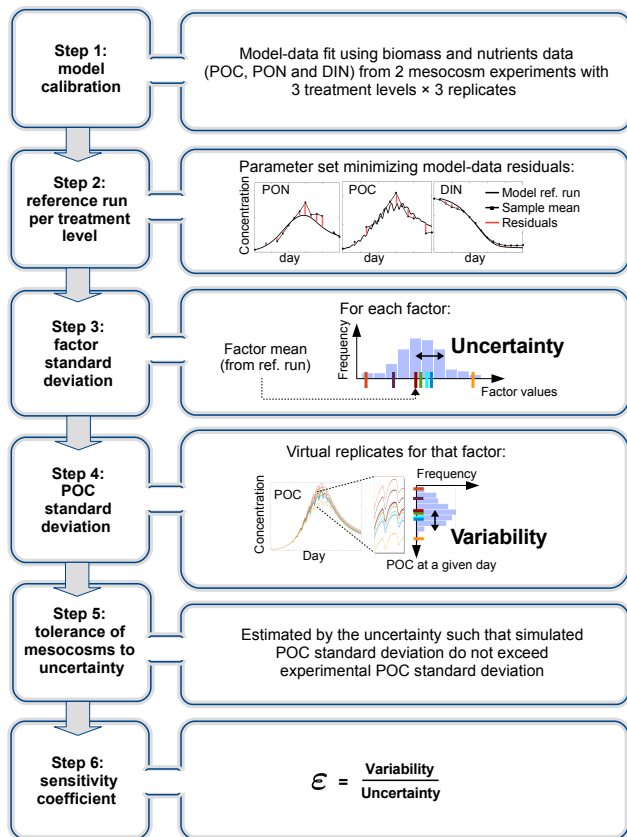
**Table 2.** Parameter values used for the reference run,  $\langle \phi_i \rangle$ . All values are common to both PeECE II and III experiments, only the mean temperature (determined by environmental forcing) and the averaged cell size in the community are different since different species composition succeeded in the experiments (*Emiliania huxleyi* was the major contributor to POC in PeECE II (Engel et al., 2008) but also diatoms significantly bloomed during PeECE III (Schulz et al., 2008)).

	Parameter	Value	Units	Variable	Reference
$a_{\text{CO}_2}$	carbon acquisition	0.15	$(\mu\text{mol-C})^{-1} \text{L}$	$\text{Phy}_C$	this study
$a_{\text{PAR}}$	light absorption	0.7	$\mu\text{mol phot}^{-1} \text{m}^2$	$\text{Phy}_C$	this study
$a^*$	carboxylation depletion	0.15	$\mu\text{m}^{-1}$	$\text{Phy}_C$	this study
$P_{\text{max}}$	max. photosyn. rate	12	$\text{d}^{-1}$	$\text{Phy}_C$	this study
$Q_{\text{subs}}^*$	subsist. quota offset	0.33	$\text{mol-N (mol-C)}^{-1}$	$\text{Phy}_C$	this study
$\alpha_Q$	$Q_{\text{subs}}$ allometry	0.4	–	$\text{Phy}_C$	this study
$\zeta$	costs of N assimil.	2	$\text{mol-C (mol-N)}^{-1}$	$\text{Phy}_C$	Raven (1980)
$\ell$	mean size $\text{Ln}(\text{ESD}/1 \mu\text{m})$	1.6	–	$\text{Phy}_C, \text{Phy}_N, \text{DIN}$	PeECE II data
		1.8	–		PeECE III data
$f_p$	fraction of protein in photosyn. machinery	0.4	–	$\text{Phy}_C, \text{Phy}_N, \text{DIN}$	this study
$V_{\text{max}}^*$	max. nutrients uptake	0.5	$\text{mol-N (mol-C d)}^{-1}$	$\text{Phy}_C, \text{Phy}_N, \text{DIN}$	this study
$\text{Aff}$	nutrient affinity	0.2	$(\mu\text{mol-C d})^{-1} \text{L}$	$\text{Phy}_C, \text{Phy}_N, \text{DIN}$	this study
$\alpha_V$	$V_{\text{max}}$ allometry	0.45	–	$\text{Phy}_C, \text{Phy}_N, \text{DIN}$	Edwards et al. (2012)
$L^*$	photosyn. losses coeff.	$11 \times 10^{-3}$	$(\mu\text{mol-C d})^{-1}$	$\text{Phy}_C, \text{Phy}_N$ and $\text{DH}_C, \text{DH}_N$	this study
$r^*$	DIN remin. & excret.	1.5	$\text{d}^{-1}$	$\text{DH}_C, \text{DH}_N$	this study
$s$	DH sinking	10	$\text{L}(\mu\text{mol-C d})^{-1}$	$\text{DH}_C, \text{DH}_N$	this study
$T_{\text{ref}}$	reference temperature	8.3	Celsius	$\text{Phy}_C, \text{Phy}_N$ and $\text{DIN}, \text{DH}_C, \text{DH}_N$	PeECE II data
		10.1	Celsius		PeECE III data

tal data are available via the data portal Pangaea (PeECE II team, 2003; PeECE III team, 2005).

Field data of aquatic  $\text{CO}_2$  concentration, temperature, and light were used as direct model inputs (see Appendix B). Measurements of POC, particulate organic nitrogen (PON), and dissolved inorganic nitrogen (DIN) were used for model calibration. Although both the experiments differ in their species composition, environmental conditions, and nutrient supply, the same parameter set was used to fit PON, POC, and DIN from PeECE II and III (i.e., a total of 54 series of repeated measures over more than two weeks), a feature indicating the model skills. In addition, the model

was validated with another 36 series of biomass and nutrients data from an independent mesocosm experiment ((Paul et al., 2014), data not shown). The experimental POC and PON data were redefined for a direct comparison with model results (see Appendix C), since some contributions (e.g., polysaccharides and transparent exopolymer particles) remain unresolved by our dynamical equations. State variables of our model comprise carbon and nitrogen contents of phytoplankton,  $\text{Phy}_C$  and  $\text{Phy}_N$ , and DIN, as representative for all nutrients. The dynamics of non-phytoplanktonic components, i.e., detritus and heterotrophs (DH), are distinguished



**Figure 1.** Variability decomposition method based on uncertainty propagation (summary of the basic principles given in Sect. 5.1.1 and 5.6.2. and Annex B in JCGM, 2008b).

by  $\text{DH}_C$  and  $\text{DH}_N$ . Thus, in our study,  $\text{POC} = \text{Phy}_C + \text{DH}_C$  and  $\text{PON} = \text{Phy}_N + \text{DH}_N$ .

The mean cell size in the community, represented as the logarithm of the mean equivalent spherical diameter (ESD), was used as a model parameter. It determines specific eco-physiological features by using allometric relations that are relevant for the computation of subsistence quota, as well as nutrient and carbon uptake rates. Regarding the latter, to resolve sensitivities to different DIC conditions, we used a relatively accurate description of carbon acquisition as a function of DIC and size. It has been suggested by previous observations and models that ambient DIC concentration increases primary production (e.g., Schluter et al., 2014; Rost et al., 2003; Zondervan et al., 2001; Riebesell et al., 2000; Chen, 1994; Riebesell et al., 1993; Riebesell and Tortell, 2011) and mean cell size in the community (Sommer et al., 2015; Eggers et al., 2014; Tortell et al., 2008). While state-of-the-art models such as Artioli et al. (2014) used empirical biomass increase to describe OA effects, we adopted and simplified a biophysically explicit description for carbon uptake from Wirtz (2011), where the efficiency of intracellular DIC transport has been derived as a function of the mean cell size  $\ell = \ln(\text{ESD}/1\mu\text{m})$  and  $\text{CO}_2$  concentration. For very large

cells, the formulation converges to the surface to volume ratio, which in our notation reads  $e^{-\ell}$ . In contrast, the dependence of primary production on  $\text{CO}_2$  vanishes for (does not apply to) picophytoplankton; the rate limitation by sub-optimal carboxylation then reads

$$f_{\text{CO}_2} = \left( \frac{1 - e^{-a_{\text{CO}_2} \cdot \text{CO}_2}}{1 + a^* \cdot e^{(\ell - a_{\text{CO}_2} \cdot \text{CO}_2)}} \right). \quad (1)$$

The specific carbon absorption coefficient  $a_{\text{CO}_2}$  reflects size-independent features of the DIC acquisition machinery (for instance, the carbon concentration mechanisms, Raven and Beardall, 2003). The coefficient  $a^*$  represents carboxylation depletion.

## 2.2 Uncertainty propagation

We considered that uncertainties were only present in the initial setup of the system; this allowed us to perform a deterministic non-intrusive forward propagation of uncertainty, which neglects the possible coupling between uncertainties and temporal dynamics unlike in intrusive methods (Chantrasmis and Iaccarino, 2012) involving stochastic dynamical equations with time-varying uncertainties (Torales and Colet, 2014; de Castro, 2017). Forward refers to the fact that unresolved differences among replicates simulated as variations of the model control factors are propagated through the model to project the overall variability in the system response, in contrast to backward methods of parameter estimation where the likelihood of input values is conditioned by the prior knowledge of the output distribution (as, for instance, in Larssen et al., 2006).

Our approach is based on a Monte Carlo method for the propagation of distributions. It is based on the repeated sampling from the distribution for possible inputs and the evaluation of the model output in each case (JCGM, 2008b). Next, the overall simulated POC variability is compared with that in POC experimental data (i.e., the mean trends of the treatment levels as well as the standard deviations are compared, the former for the calculation of the reference run and the latter for the uncertainty propagation). Among the available experimental data, we favored POC over PON and DIN in the uncertainty propagation analysis since it is usually the target variable of OA effects and shows the highest variability. A variability decomposition with more than one dependent variable (equivalent to a multivariate ANOVA design, for instance) is beyond of the scope of the study. The comparison between simulated and experimental variability in POC helps in the identification of the changes in physiological state and community structure that are the main potential contributors to the variability.

We considered model factors,  $\phi_i$ , with  $i = 1, \dots, N = 19$ , consisting of 14 process parameters and 5 initial conditions for the state variables. Their reference values,  $\langle \phi_i \rangle$ , were adjusted to yield model solutions reproducing the mean of

each treatment level (steps 1 and 2, Tables 1 and 2). To test our first hypothesis, factor variations representing potential uncertainties are introduced as random values distributed around  $\langle\phi_i\rangle$  with standard deviation  $\Delta\phi_i$ . To calculate  $\Delta\phi_i$ , we first generate  $10^4$  simulations, each one with a different factor value,  $\phi_i$  (steps 3 and 4). The ensemble of model solutions for each factor and treatment level simulates the potential experimental outcomes, hereafter referred to as “virtual replicates”, (see Appendix D). The factor value for each POC trajectory is randomly drawn from a normal distribution around the factor reference value  $\langle\phi_i\rangle$  (same distribution is assumed by popular parametric statistical inference tools such as regressions and ANOVA, Field et al., 2008). For every treatment level and at every time step, we calculated the ensemble average of the virtual replicates,  $\langle\text{POC}_i^{\text{mod}}\rangle$ , and the standard deviation,  $\Delta\text{POC}_i^{\text{mod}}$ . Thus,  $\Delta\phi_i$  is the standard deviation of the distribution of factor values, such as  $\Delta\text{POC}_i^{\text{mod}}$ , which do not exceed the standard deviation of the experimental POC data,  $\Delta\text{POC}^{\text{exp}}$ , for any mesocosm at any given time (step 5). The effect of variations of  $\phi_i$  on the variability (step 6) is given as follows:

$$\varepsilon_i = \frac{\Delta\text{POC}_i^{\text{mod}}}{\Delta\phi_i}. \quad (2)$$

This ratio expresses the maximum variability a factor can generate,  $\Delta\text{POC}_i^{\text{mod}}$ , relative to the associated range of that factor variations,  $\Delta\phi_i$ , to ensure that  $\Delta\text{POC}_i^{\text{mod}}$  is the closest to  $\Delta\text{POC}^{\text{exp}}$  at any time. In general,  $\varepsilon_i$  defines how much of the uncertainty of a dependent variable  $Y$  (here  $Y = \text{POC}$ ) is explained by and the uncertainty of the input factors  $\phi_i$ , a proxy of which is known as the sensitivity coefficient  $c_i = \frac{\partial Y}{\partial \phi_i}$  in the widespread formula to calculate error propagation (Ellison and Williams, 2012), also known as law of propagation of uncertainty (JCGM, 2008a)

$$(\Delta Y)^2 = \sum_{i=1}^N c_i^2 \cdot (\Delta\phi_i)^2. \quad (3)$$

This expression is based on the assumption that changes in  $Y$  in response to variations in one factor  $\phi_i$  are independent from those owing to changes in another factor  $\phi_j$ , and that all changes are small (thus cross-terms and higher-order derivatives are neglected). Where no reliable mathematical description of the relationship  $Y(\phi_i)$  exists (in our case, only an expression for the rate equation  $d\text{POC}/dt$  is known (see Table 1) but not its analytical solution, i.e., POC),  $c_i$  can be evaluated experimentally (Ellison and Williams, 2012; JCGM, 2008a). As mentioned in the Introduction and Appendix A, such high-dimensional multi-factorial measurements are costly in mesocosm experiments. Therefore, we obtained equivalent information by numerically calculating  $\varepsilon_i$ . Such approximations to sensitivity coefficients calculated by our Monte Carlo method of uncertainty propagation correspond to taking all higher-order terms in the Taylor series expansion into account since no linearization is required

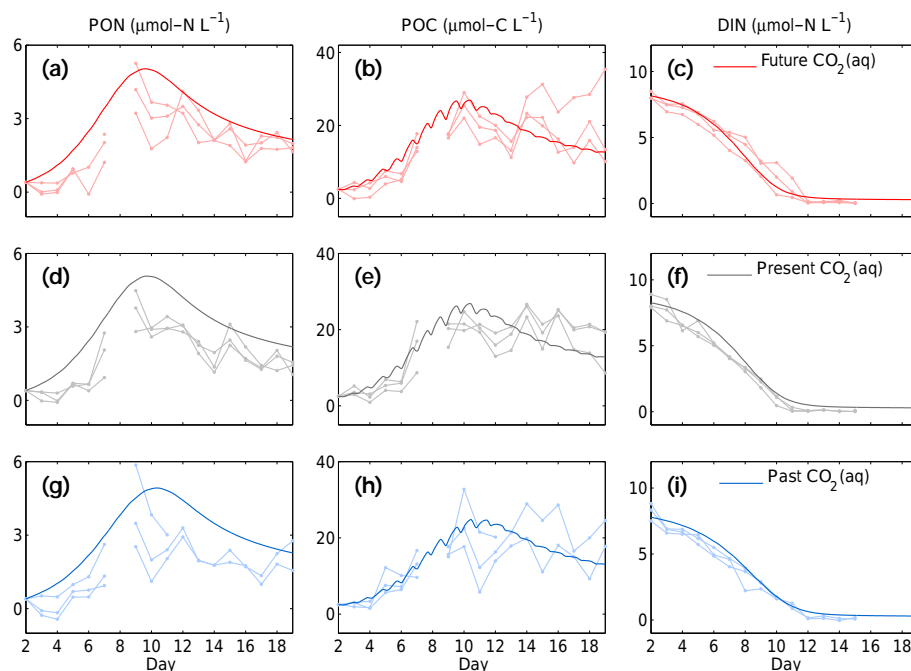
(see Sect. 5.10 and 5.11 and Annex B in JCGM, 2008b). A straightforward extension including the cross-terms showing synergistic uncertainties effects, as in an experimental multi-way ANOVA design, requires the assumption of joint distributions for the uncertainty of factors and the calculation of covariance matrices, a considerable effort that is beyond of the scope of this paper.

Hereafter, the standard deviation of any given factor, i.e., factor uncertainty, will be given as percentage of the reference values and will be called  $\Delta\Phi_i$ . The actual factor range is given as  $\Delta\phi_i = \frac{\Delta\Phi_i \cdot \phi_i}{100}$ . Strong irregularities in the standard deviations of experimental POC data (for instance, small  $\Delta\text{POC}^{\text{exp}}$  at day 8 in Fig. 2p), translates to remarkably enhanced or reduced sensitivity coefficients if the model–data comparison would be performed at a daily basis. Therefore, we considered the temporal mean of the standard deviation per phase, i.e., prebloom, bloom, and postbloom. We inferred phases for PeECE II from Engel et al. (2008) and for PeECE III from Schulz et al. (2008) and Tanaka et al. (2008).

To numerically calculate the ensemble of  $10^4$  POC trajectories per factor (i.e., the virtual replicates; see Fig. 8), we applied the Heun integration method with a time step of  $4 \times 10^{-4}$ , (about 35 s of experimental time). The number of simulated POC time series is chosen such as a higher number of model realizations, i.e., a higher number of virtual replicates will produce the same results (see Adaptive Monte Carlo procedure, Sect. 7.9. in JCGM, 2008b). We dismissed the negative values that randomly appeared when drawing  $10^4$  values from the normal distribution of factor values; this reduction in the number of trajectories did not affect the results.

Environmental data showed low variability among similar treatment replicates, (see Fig. 9), suggesting a non-direct relation between variations in environmental factors among replicates and the observed biomass variability. Therefore, we focused on uncertainties in ecophysiology and community composition and used environmental data as forcing. Perturbations of the similarity among replicates produced by strong changes in environmental conditions (storms, dysfunctional devices, etc.) or by errors in manipulation or sampling procedures are not within the scope of this work. After a few decades, the current state-of-the-art of experimental techniques for running plankton mesocosms is advanced. We believe such differences are of low impact or well understood in terms of their consequences for final outcomes (Riebesell et al., 2010; Cornwall and Hurd, 2015).

Notably, our analysis suggested sufficient (but not necessary; Brennan, 2012) causes of uncertainties in mesocosm experiments. Variations in model characterization including structural variability (Adamson and Morozov, 2014; Fussmann and Blasius, 2005) or uncertainties in model parameterization (Kennedy and O’Hagan, 2001) or comparisons to different uncertainty propagation methods (de Castro, 2017) require further extensive analyses, which is beyond the scope of this study. However, we performed a series of preceding



**Figure 2.** Solid lines show reference runs for POC, PON, and DIN simulating the mean of the replicates per treatment level, with different colors for the three experimental CO<sub>2</sub> setups. Dots are replicated data from the Pelagic Enrichment CO<sub>2</sub> Experiment (PeECE II) for newly produced POC and PON, i.e., starting values at day 2 were subtracted from subsequent measurements as in Riebesell et al. (2007).

model analyses (including uncertainty propagation) by using slightly different model formulations (data not shown). From these preceding analyses, we found that different model formulations can lead to quantitatively different confidence intervals, but leave the final results qualitatively unchanged.

### 3 Results

#### 3.1 CO<sub>2</sub> effect on POC dynamics

Our model reproduces the means of PON, POC, and DIN experimental data per treatment level, i.e., for the future, present, and past CO<sub>2</sub> conditions, in two independent PeECE experiments (Figs. 2 and 3). For PeECE II, PON is moderately overestimated and postbloom POC is slightly underestimated. Nonetheless, the model represents the experimental data with similar precision than the means of experimental replicates (see Appendix E). The means of the same treatment replicates and their associated standard deviations are typically used to represent experimental data (see Fig. 1b in Engel et al., 2008 for PeECE II or Fig. 8a in Schulz et al., 2008 for PeECE III). The means are in the foundations of the statistical inference tools that did not detect acidification responses for PeECE II and III. However, with our mechanistic model-based analysis, phytoplankton growth in the future CO<sub>2</sub> conditions showed an earlier and elevated bloom with respect to past CO<sub>2</sub> conditions. The future and past reference trajectories limit the range of the CO<sub>2</sub> enrichment effect, as

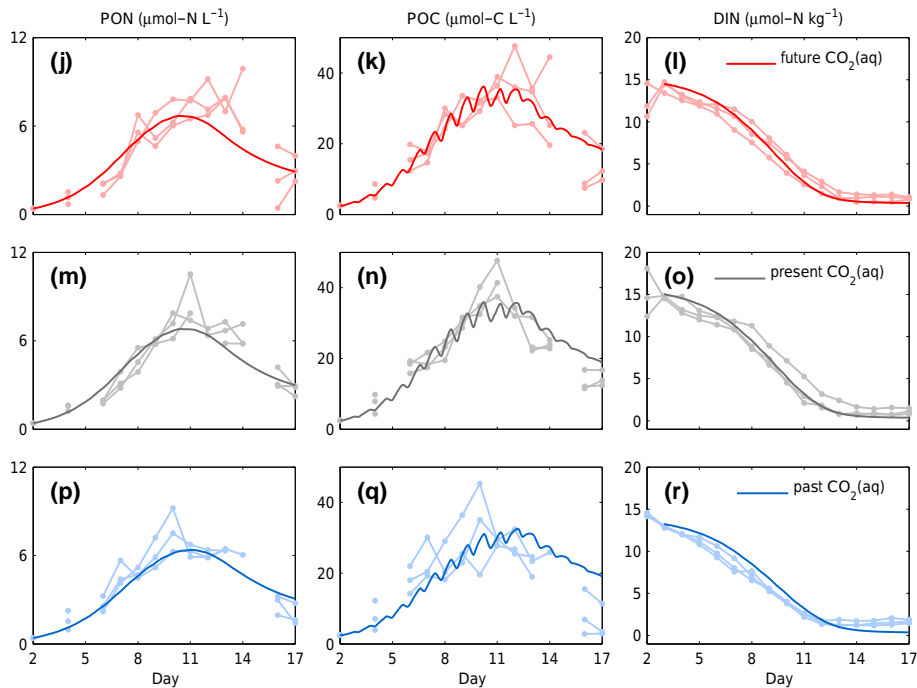
shown by the dark gray area in Fig. 4. POC variability owing to variations in model factors simulating experimental uncertainties is plotted as the light gray area in the figure. Our results suggest that avoiding high POC standard deviations that potentially mask OA effects in experimental data requires the reduction of the factor variations triggering variability during the bloom.

#### 3.2 CO<sub>2</sub> effect on uncertainty propagation

The estimation of the tolerance thresholds of the dynamics to uncertainty propagation for the two test-case experiments, per acidification levels and per factor uncertainty, are listed in Table 3. We investigated the potential interaction of the treatment and the uncertainty effects on the tolerance by a linear mixed-effects model with  $\phi_i$  as the random factor (R Core Team, 2016). The synergistic effect between the factor uncertainty and the treatment levels was found to be non-significant ( $F = 2.9$  with  $p = 0.06$ ). Therefore, the thresholds do not appear to statistically depend on the treatment level, even when the standard deviation of the measured POC data,  $\Delta\text{POC}^{\text{exp}}$ , for the future and past acidification conditions were, on average, about 70 % larger than the standard deviation of the present conditions (POC experimental data in Figs. 2 and 3 are more spread in the future and past concentrations than in the present concentration). Despite the low statistical power of this test (only data from two independent samples, the two PeECE experiments, were available),

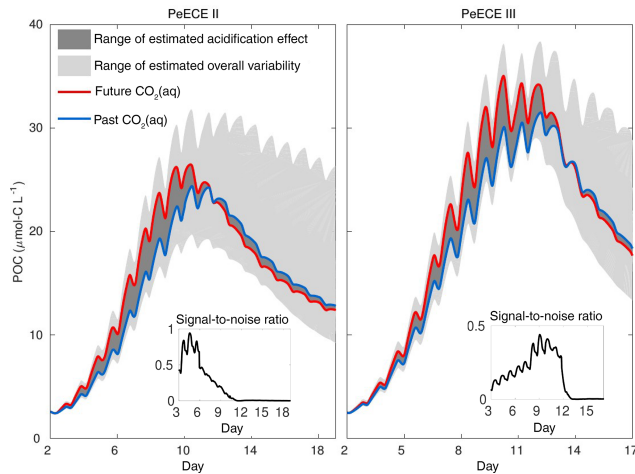
**Table 3.** Tolerance of mesocosms experiments to differences among replicates, given as a percentage of the reference factor value listed in Tables 1 and 2. According to our model projections, above these thresholds the simulated variability,  $\Delta\text{POC}_i^{\text{mod}}$ , exceeds the observed variability,  $\Delta\text{POC}_i^{\text{exp}}$ . Main contributors to the simulated variability during the bloom are highlighted in bold (see Sect. 3).

Factor $\phi_i$		$\Delta\Phi_i$ (%)						Averaged tolerance
		PeECE II			PeECE III			
		Future	Present	Past	Future	Present	Past	
								(%)
Phy <sub>C</sub> (0)	initial phyto C biomass	68	49	46	78	60	100	67 ± 6
Phy <sub>N</sub> (0)	initial phyto N biomass	26	19	22	21	16	29	22 ± 4
<b>DIN(0)</b>	<b>initial DIN</b>	<b>20</b>	<b>28</b>	<b>29</b>	<b>17</b>	<b>11</b>	<b>18</b>	<b>20±6</b>
$a_{\text{CO}_2}$	carbon acquisition	89	46	23	86	63	46	59 ± 23
$a_{\text{PAR}}$	light absorption	>100	>100	98	>100	>100	92	> 100
$P_{\text{max}}$	maximum photosyn. rate	27	18	16	22	16	28	21 ± 5
$Q_{\text{subs}}^*$	subsistence quota offset	6	5	6	5	4	9	6 ± 1
$\alpha_Q$	$Q_{\text{subs}}$ allometry	9	7	8	7	5	10	8 ± 2
$\ell$	<b>size Ln(ESD/1µm)</b>	<b>25</b>	<b>20</b>	<b>29</b>	<b>19</b>	<b>14</b>	<b>22</b>	<b>22±5</b>
$f_p$	fraction of protein in photosyn. machinery	92	75	44	36	17	38	50 ± 25
$V_{\text{max}}^*$	maximum nutrient uptake	13	11	14	10	8	14	12 ± 2
Aff	nutrients affinity	39	31	42	38	36	55	40 ± 7
$\alpha_V$	$V_{\text{max}}$ allometry	14	11	15	10	8	14	12 ± 2
$L^*$	<b>phytoplankton losses</b>	<b>22</b>	<b>30</b>	<b>28</b>	<b>12</b>	<b>10</b>	<b>15</b>	<b>20±8</b>
$r^*$	DIN remineralization	73	99	98	128	37	52	81 ± 31
$s$	DH sinking	> 100	> 100	96	> 100	61	79	>100
$T_{\text{ref}}$	reference temperature	17	12	14	9	7	14	12 ± 3



**Figure 3.** As in Fig. 2 for PeECE III.





**Figure 4.** Reference simulations of POC for high  $\text{CO}_2$  (red) and low  $\text{CO}_2$  (blue) conditions bind the range of acidification effect (dark gray) according to our model projections. Light gray area shows the limits of the overall simulated POC variability,  $\Delta\text{POC}^{\text{mod}}$ . Inlay graph display the signal-to-noise ration (black solid lines), i.e., the ratio between the variance of the acidification effect and the variance of the overall variability.

we still considered the potential lack of  $\text{CO}_2$  effect on the uncertainty propagation as sufficient justification to simplify further analysis on variability decomposition by averaging the thresholds and the sensitivity coefficients over treatment levels (see last column in Table 3 and Fig. 7).

### 3.3 Variability decomposition

Our method allows for decomposition of POC variability in factor-specific components  $\Delta\text{POC}_i^{\text{mod}}$ . The effect of factor variations simulating experimental differences among replicates is classified depending on its nature, intensity, and timing (Figs. 5, 6, and 7).

POC variability during the prebloom phase can be explained mainly by the differences of factors related to subsistence quota, i.e.,  $Q_{\text{subs}}^*$  and  $\alpha_Q$ , in both PeECE II and III experiments (left column in Figs. 5 and 6). This suggests that the differences in subsistence quota first intensify the divergence of POC trajectories and then weaken a few days later because of the system dynamics. These subsistence parameters only need to vary about 6 and 8 % among replicates (see Table 3) to maximize their contribution to the  $\Delta\text{POC}^{\text{exp}}$ ; thus, their sensitivity coefficients are the highest (see Fig. 7).

Differences in the initial nutrient concentration,  $\text{DIN}(0)$ , mean cell size,  $\ell$ , and phytoplankton biomass loss coefficient,  $L^*$ , generate the modeled variability mainly during the bloom (with just about 20 % differences among replicates; see Table 3 and second column in Fig. 5) showing high values of the sensitivity coefficient (highlighted in Fig. 7).

Amplified variability in the postbloom phase (third column in Figs. 5 and 6) can be attributed to the uncertainties

in the reference temperature  $T_{\text{ref}}$  for the Arrhenius equation, Eq. (A4), in sinking loss or export flux,  $s$ , and in remineralization and excretion,  $r^*$ . The sensitivity coefficient of  $T_{\text{ref}}$  is high, with just about 12 % variation. Therefore, even if differences in ambient temperature among replicates of the same sample are negligible (see the low standard deviations in the temperature, Fig. 9), differences in the metabolic dependence on that ambient temperature seems to be relevant in the decay phase. Interestingly, variations among replicates in the physiological dependence on other environmental factors do not show the same relevance (the sensitivity coefficient  $\varepsilon_i$  is low for carbon acquisition  $a_{\text{CO}_2}$  and light absorption  $a_{\text{PAR}}$ ). Generating high divergence during the postbloom requires a strong perturbation of parameters for the description of the non-phytoplanktonic biomass (about 81 % of the reference value for sinking and 96 % for remineralization and excretion, see Table 3), which translates to a relatively low sensitivity coefficient.

Perturbations of the initial detritus concentration,  $\text{DH}_C(0)$  and  $\text{DH}_N(0)$  have no impact on the dynamics provided that they remain within reasonable ranges ( $\Delta\Phi_i < 100$ ). In fact, more than 10-fold difference among replicates in such non-relevant factors were necessary to achieve a perceptible variability  $\Delta\text{POC}_i^{\text{mod}}$ .

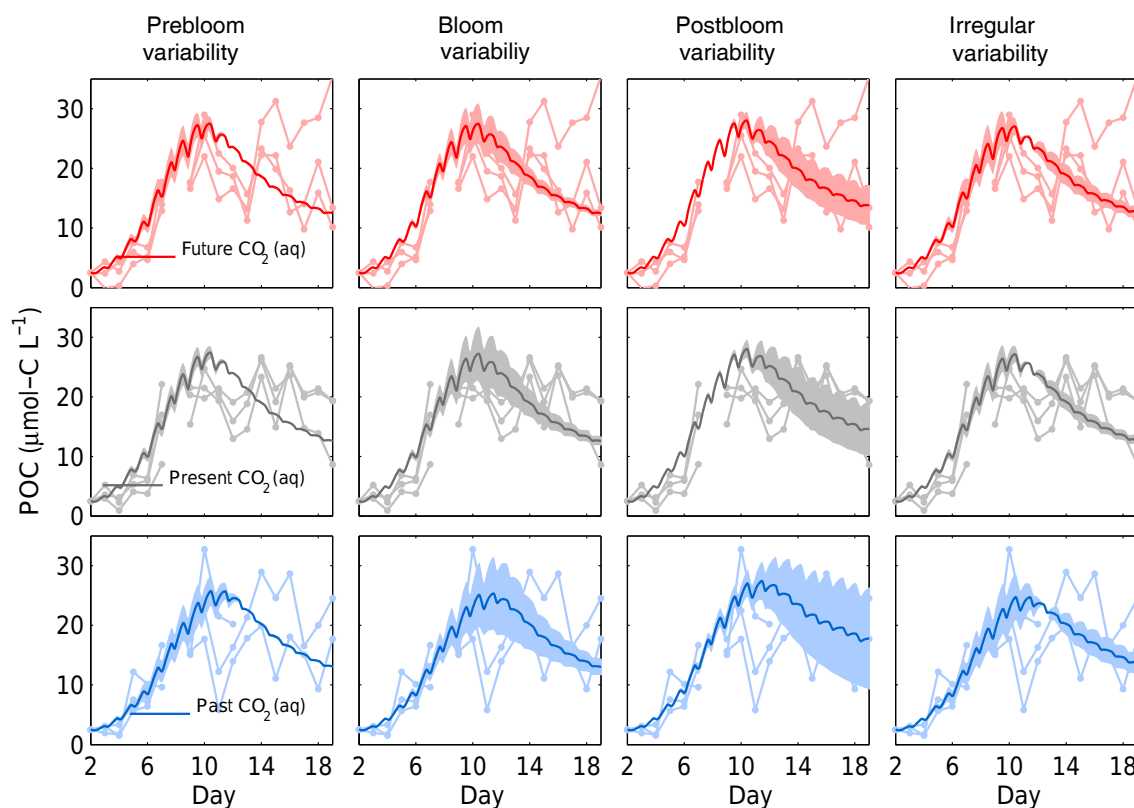
POC variability throughout the bloom phases (right column in Figs. 5 and 6) can be attributed to the varying carbon and nitrogen initial conditions,  $\text{Phy}_C$  and  $\text{Phy}_N$ , nutrient uptake-related factors,  $V_{\text{max}}^*$ ,  $\alpha_V$ , and  $\text{Aff}$ , and protein allocation for photosynthetic machinery,  $f_p$ . With regard to the latter, high standard deviations of the tolerance (see Table 3) suggest non-conclusive results.

## 4 Discussion

We used the uncertainty quantification method to decompose POC variability by using a low-complexity model that describes the major features of phytoplankton growth dynamics. The model fits the mean of mesocosm experimental PeECE II and III data with high accuracy for all  $\text{CO}_2$  treatment levels. We confirmed the working hypotheses (Figs. 5–7); in particular, we showed that small differences in initial nutrient concentration, mean cell size, and phytoplankton biomass losses are sufficient to generate the experimentally observed bloom variability  $\Delta\text{POC}^{\text{exp}}$  that potentially mask acidification effects, as discussed in the following subsections.

The results of our analyses are conditioned by the dynamical model equations imposed. Deliberately, the model's complexity is kept low, mainly to limit the generation of structural errors with respect to model design. At the same time, the level of complexity resolved by the model suffices to explain POC measurements of two independent mesocosm experiments with identical parameter values (see Table 2), which highlights model skill. The used equations comply





**Figure 5.** POC variability decomposition per factor,  $\Delta\text{POC}_i^{\text{mod}}$  for PeECE II. Shaded areas are limited by the standard deviation of  $10^4$  simulated POC time series (see Sect. 2), around the mean trajectory of the ensemble (solid line). The timing of the amplification of the variability determines four separated kinds of behavior: factor uncertainties generating variability during the prebloom, bloom, postbloom, or at irregular phase (see Sect. 3).

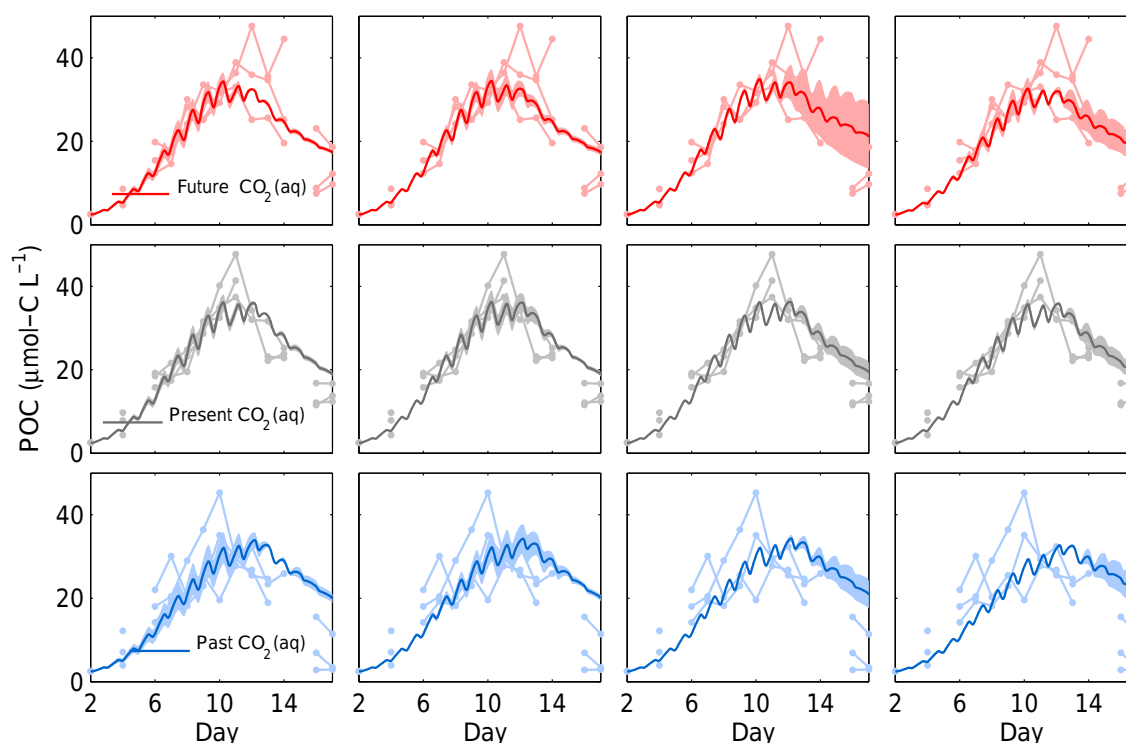
with theories of phytoplankton growth (e.g., Droop, 1973; Aksnes and Egge, 1991; Pahlow, 2005; Edwards et al., 2012; Litchman et al., 2007; Wirtz, 2011). The uncertainty propagation employed here can be applied to any model. As long as the model features a similar structural complexity and is also able to reproduce POC with sufficient accuracy, we expect similar qualitative findings with respect to the factors ( $\Phi_i$ ) and similar identification of the major contributors to the variability. However, we would not expect other models to reveal exactly similar values in the ratio  $\epsilon_i$ , which would likely depend on the equations used to resolve some of the ecophysiological details.

#### 4.1 Nutrient concentration

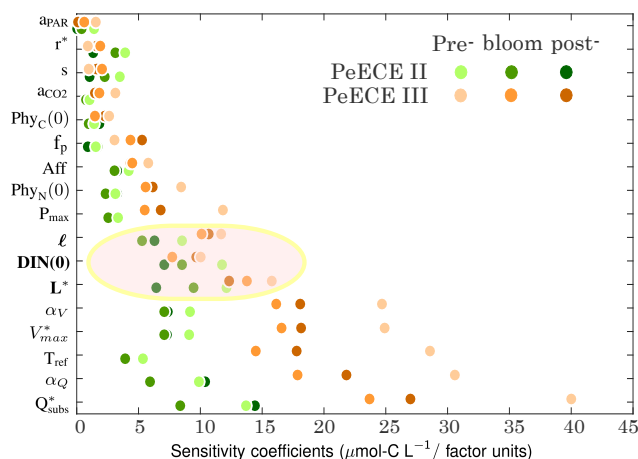
Differences among replicates in the initial nutrient concentration substantially contribute to POC variability, a sensitivity that is, interestingly, not well expressed when varying the initial cellular carbon or nitrogen content of the algae,  $\text{Phy}_C(0)$  and  $\text{Phy}_N(0)$ . The relevance of accuracy for the initial nutrient concentration in replicated mesocosms has already been pointed out in Riebesell et al. (2008). Under a constant growth rate,  $\text{DIN}(0)$  determines the timing of nu-

trient depletion; therefore, differences in the initial nutrient concentration might also translate into temporal variations in the succession of species. We showed that such dependence is noted even in more general dynamics, and that our method can also estimate the variational range for differences in the initial DIN concentration for experiments with a low number of replicates. The standard deviation of  $\text{DIN}(0)$  in the experimental setup for PeECE III was 50 % of the mean, which is significantly above our tolerance threshold (see Table 3 for initial DIN concentration). Following Riebesell et al. (2007), we considered day 2 as the initial condition, when the measured DIN was  $14 \pm 2 \mu\text{mol-CL}^{-1}$ , as shown in Table 1. Since  $2 \mu\text{mol-CL}^{-1}$  is approximately 14 % of  $14 \mu\text{mol-CL}^{-1}$ , the variability of replicates at day 2 was about 14 %. Therefore, experimental differences in the initial nutrient concentration were similar to the tolerance threshold for the initial DIN established to avoid high variability ( $(20 \pm 6) \%$  in Table 3), which represents an explanation for the high divergence observed in POC measurements.

For PeECE II, experimentally measured DIN concentration at day 0 was  $10.7 \pm 0.8 \mu\text{mol-CL}^{-1}$ , suggesting a 7.5 % difference among replicates, which was below our projected tolerance level (7.5 is out of the range [14, 26]). The same



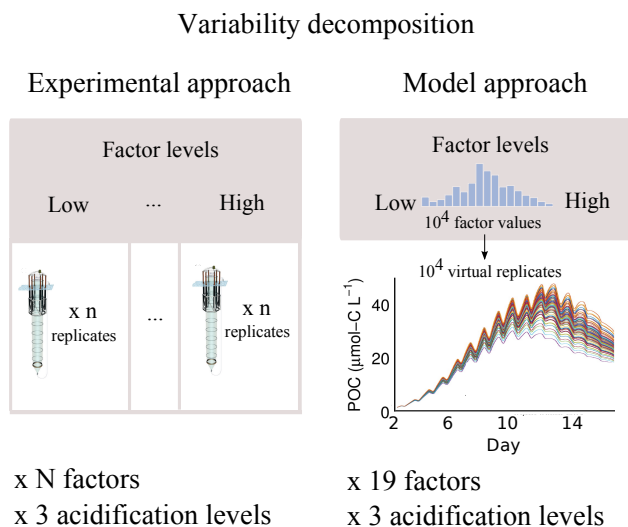
**Figure 6.** As Fig. 5, for PeECE III.



**Figure 7.** Sensitivity coefficients ( $\varepsilon_i$ ; Eq. 2) of factors  $\phi_i$  listed in Tables 1 and 2 for different bloom phases in two OA-independent mesocosm experiments. Factors whose uncertainties potentially mask acidification effects (Fig. 4) by triggering variability during the bloom (Figs. 5 and 6) are highlighted.

was noted for day 2, with DIN concentration equal to  $8 \pm 0.5 \mu\text{mol-C L}^{-1}$  (Table 1). Our approach showed that differences in initial nutrient concentration in PeECE II were not sufficiently high to trigger the experimentally observed POC variability. Incidentally, phosphate re-addition on day 8 of the experiment established new initial nutrient concen-

tration for the subsequent days. When the dynamics in one replicate significantly diverges from the mean dynamics of the treatment, even if the re-addition occurs at the same time and at the same concentration in all the replicates, the mesocosm with the outlier trajectory will not respond as the others do, and with the addition of a new nutrient condition, the divergence might be further amplified. In this case, nutrient re-addition has the same impact on the systems as variations in the initial conditions of nutrient concentration. Also for PeECE II, variability in POC is about 30% higher than that for PON, as shown in Fig. 2. We attribute the temporal decoupling between C and N dynamics to the break of symmetry among replicates by the nutrient re-addition, owing to the strong sensitivity of the system to initial nutrient concentrations and a concomitant change in subsistence N : C quota, which is a sensitive parameter, especially during the pre-bloom phase (Figs. 5, 6, and 7). Increase of POC : PON ratios under nitrogen deficiency has been observed frequently during experimental studies (e.g., Antia et al., 1963; Biddanda and Benner, 1997) and has been attributed to preferential PON degradation and to intracellular decrease of the N : C ratio (Schartau et al., 2007). Hence, we confirmed that nutrient re-addition during the course of the experiments results in a significant disturbance, as has been previously reported (Riebesell et al., 2008), although a complete analysis would require a model that explicitly accounts for other nutrients.



**Figure 8.** The exploration of the sources of variability in an experiment with a multi-way repeated measures ANOVA design with 3 acidification levels requires a multi-factorial high-dimensional setup (left panel). Alternately, we numerically simulate the biomass dynamics with  $10^4$  virtual replicates, each one with a different normally distributed factor value (right panel). Uncertainty propagation relates the dispersion of the factor values with the dispersion of the POC trajectories. As an example, we plot results of POC variability in 50 virtual replicates of PeECE III at low acidification with uncertainty in initial nutrient concentration. Mesocosm drawing from Scheinin et al. (2015).

#### 4.2 Mean cell size as a proxy for community structure

We found a limited tolerance to variations in the mean cell size of the community,  $\ell$ , which has a threshold of about 22 % variation (see Table 3). If we consider the averaged mean cell size of PeECE II,  $\langle \ell \rangle = 1.6$ , and III,  $\langle \ell \rangle = 1.8$ , from Table 2, we obtain  $\langle \ell \rangle = 1.7$ . Then the absolute standard deviation is  $\Delta \ell = 22 \cdot \frac{1.7}{100} \sim 0.4$ . Therefore, our methodology shows that variations within the range limited by  $\langle \ell \rangle \pm \Delta \ell$ , i.e., [1.3, 2.1], are sufficient to reproduce the observed experimental POC variability during the bloom. Since  $\ell$  is in the log scale, the corresponding ESD increment is within the variational range  $\langle \text{ESD} \rangle \pm \Delta \text{ESD}$ , that is, [3.7, 8.1]  $\mu\text{m}$  (or [25, 285]  $\mu\text{m}^3$  in volume). These values are easily reached in the course of species succession. This supports studies showing that community composition outweighs ocean acidification (Eggers et al., 2014; Kroeker et al., 2013; Kim et al., 2006).

#### 4.3 Phytoplankton loss

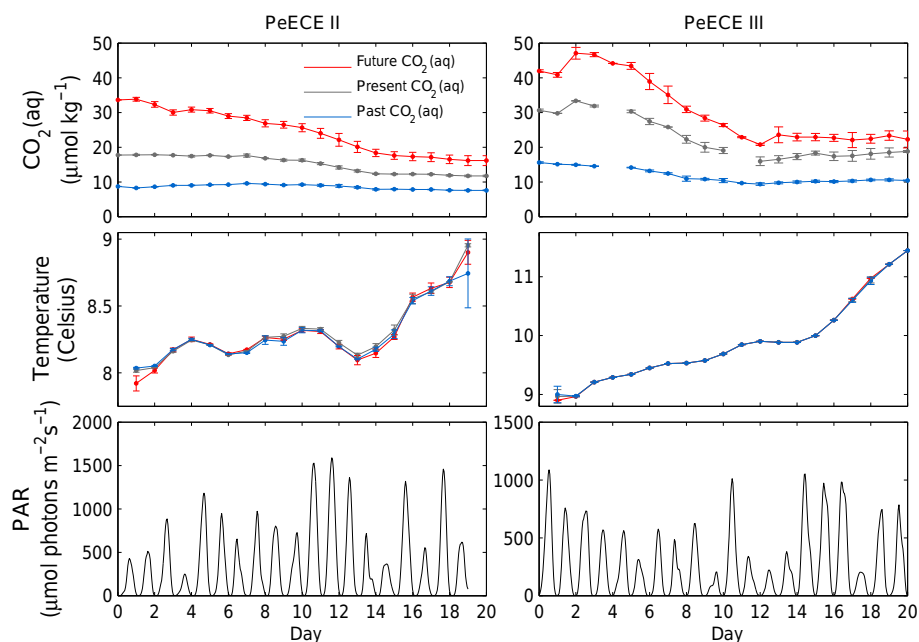
Another major contributor to POC variability during the bloom phase is phytoplankton biomass loss,  $L^*$ . With a standard deviation of about 20 % (Table 3), uncertainties in  $L^*$  generate variability larger than the model response to OA, in particular at the end of the growth phase and the beginning

of the decay phase. Unresolved details in phytoplankton loss rate include, among others, replicate differences in cell aggregation or damage by collisions, mortality by virus, parasites, and morphologic malformations, or grazing by non-filtered mixotrophs or micro-zooplankton.

#### 4.4 Inference from summary statistics on mesocosm data with low number of replicates

To test the hypotheses outlined in the Introduction entails two important aspects. First, heuristic exploration of variability would require experiments designed to quantify the sensitivity of mesocosms to variations in potentially relevant factors that specify uncertainties in environmental conditions, cell physiology, and community structure. However, this would require high-dimensional multi-factorial setups (see Appendix D), which would be difficult to handle, if at all, even for low number of replicates. Second, standard statistical inference tools might come to their limitations in estimating treatment effects. Repeated measures of relevant eco-physiological data (e.g., POC) are collected from mesocosm experiments that span a few weeks. If the differences among treatment levels are smaller than those among replicates of the same treatment level, post-processing statistical analyses might conclude that there are no detectable effects (Field et al., 2008).

In many cases, the mean and the variance of the sample are taken as a fair statistical representation of the effect of the treatment level and its variability. However, summary statistics such as the mean and the variance might fail to describe distributions that do not cluster around a central value, i.e., when the data are not normally distributed in the sample. This is because a feature of normally distributed ensembles is that the mean represents the most typical value and deviations from that main trend (caused by unresolved factors not directly related to the treatment) might cancel out in the calculation of the ensemble average. Actually, this cancellation is the reason for using replicates (Ruxton and Colegrave, 2006), but many circumstances can remarkably lower the likelihood for cancellation, for instance, (i) effects that are sensitive to initial conditions (thus, small initial differences in the replicates of a given sample might become amplified and produce departures that enlarge over the course of the experiment), (ii) non-symmetrically distributed initial conditions in the sample (that might lead to non-symmetrical distribution of the results), and (iii) a low number of replicates, i.e., a sample size not adapted to the intensity of the treatment effect, the sensitivity of all effects to initial conditions, and the intended accuracy of the experiment. Each incident decreases the statistical power and therefore misleading conclusions might be inferred (Miller, 1988; Cohen, 1988; Peterman, 1990; Cottingham et al., 2005).



**Figure 9.** Environmental data from PeECE II and III are taken as model inputs. Error bars denote the standard deviation of the same treatment replicates.

#### 4.5 Consequences for the design of mesocosm experiments

In our simulations, the  $\text{CO}_2$  level affected the intensity and timing of the bloom (Fig. 4). Thus, the slope of the growth phase can be regarded as a suitable target variable to detect OA effects. Moreover, our model analysis revealed a low signal-to-noise ratio. The ability to distinguish the treatment effect from noise depends on the experimental design, the strength of the treatment, and the variability that it is not explained by the treatment. When the signal-to-noise ratio is as low as it is shown by our simulations, a large experimental sample size is needed to avoid incurring a type II error (Field et al., 2008). In particular, we can assume a two sample two-sided balanced  $t$  test with two treatment levels as in Fig. 4, i.e., the maximum difference between means equal to approximately  $5 \mu\text{mol-CL}^{-1}$  (see, i.e., PeECE III at day 10) and the variability  $\Delta\text{POC}^{\text{mod}}$  approximately equal to  $4 \mu\text{mol-CL}^{-1}$ . If we aim for a statistical power of 0.8, i.e., a 80 % chance of observing a statistically significant result with that experimental design, the required number of replicates per treatment level would be 11 (R Core Team, 2016), which is unpractical in mesocosm experiments. With  $n = 3$  replicates, the chance declines to only 20 %.

We provided an estimation for the uncertainty thresholds that can be used for improving future sampling strategies with a low number of replicates, i.e.,  $n = 3$ . Tolerances shown in Table 3 can be used to quantify how much replicates similarity can be compromised before the variability of the outcomes outweighs potential acidification effects. Some

tolerances indicate maximal variations in observable quantities, such as nutrient concentration and community composition. These model results suggest that a better control of such dissimilarities among replicates can help maintain the variability below the range of the acidification effect, especially during the bloom.

Strategies to reduce  $\Delta\text{POC}^{\text{mod}}$  should similarly apply to lower  $\Delta\text{POC}^{\text{exp}}$ . For example, model results turned out to be very sensitive to variations in mean logarithmic cell size. Variations of this factor during the initial filling of the mesocosms may already generate divergent responses in POC so that a potential  $\text{CO}_2$  signal becomes difficult to detect, if at all. To determine spectra of cell sizes (or mean of logarithmic cell size) of the initial plankton community prior to  $\text{CO}_2$  perturbation would be a possibility to countervail this difficulty. The decision of which mesocosm to select for which kind (i.e., intensity) of perturbation may then be adjusted according to similarities in initial plankton community structure. For example, we may consider some number of available mesocosms that should become subject to two different  $\text{CO}_2$  perturbation levels. We may first select two mesocosms that reveal the greatest similarity with respect to their initial size spectra and assign them to the two different  $\text{CO}_2$  treatment levels. Likewise, from the remaining mesocosms we again chose those two mesocosms that show the closest similarity between their size spectra. Those two are chosen to become subject to the two different  $\text{CO}_2$  perturbations. The selection procedure could be repeated until all mesocosms have been assigned to either of the two  $\text{CO}_2$  treatments. Thus, mesocosms with similar initial conditions are assured to become

subjected to different CO<sub>2</sub> perturbations. This reduces a mixture of random effects due to variations in experiment initialization and CO<sub>2</sub> effect and it will likely facilitate data analysis in experimental setups with low number of replicates, where sample randomization (Ruxton and Colegrave, 2006) might not be effective; see Sect. 4.4. Mesocosms may then be first analyzed pairwise (similar initial setup) with respect to differences in CO<sub>2</sub> response.

In addition, our analysis results help interpreting non-conclusive results and provide plausible explanations for the negative results for the detection of potential acidification effects (Paul et al., 2015; Schulz et al., 2008; Engel et al., 2008; Kim et al., 2006; Engel et al., 2005). Thus, our study also suggests the limitation of the statistical inference tools commonly used to assess the statistical significance of effect detectability.

Finally, we found the same main contributors to POC variability for all the treatment levels, even if experimental variability is about 70% higher in the mesocosms where the carbon chemistry was manipulated. In particular, the heterogeneity of variance measured in future levels is larger than under the other acidification conditions (see fluctuations of the standard deviations of CO<sub>2</sub> concentrations, Fig. 9). These differences in biomass variability among treatment levels are not explained by uncertainties in our model factors. They might have been originated by the irregularities in the CO<sub>2</sub> aeration (Riebesell et al., 2008; Cornwall and Hurd, 2015); however, further analyses need to be conducted to determine potential sources of *differences* in variability.

## 5 Conclusions

Our model projections indicated that phytoplankton responses to OA were mainly expected to occur during the bloom phase, presenting a higher and earlier bloom under acidification conditions. Moreover, we found that amplified POC variability during the bloom that potentially reduces the low signal-to-noise ratio can be explained by small variations in the initial DIN concentration, mean cell size, and phytoplankton loss rate.

The results of the model-based analysis can be used for refinements of experimental design and sampling strategies. We identified specific ecophysiological factors that need to be confined in order to ensure that acidification responses do not become masked by variability in POC.

With our approach we reverse the question of how experimental data can constrain model parameter estimates and instead determine the range of variability in experimental data that can be explained by modeling with variational ranges bounding uncertainties of specific control factors. We tested the hypothesis of whether small differences among replicates have the potential to generate higher variability in biomass time series than the response that can be attributed to the effect of CO<sub>2</sub>. Therefore, we conclude that modeling studies that integrate data from acidification experiments should resolve physiological regulation capacities at cellular and community levels. In fact, modeling the propagation of uncertainties revealed cell size to be a major contributor to phytoplankton biomass variability. This suggests the use of adaptive size-trait-based dynamics since such approaches allow for the resolution of ecophysiological trait shifts in non-stationary scenarios (Wirtz, 2013). The role of intracellular protein allocation can also be clarified by using a trait-based approach, since our results about the impact of its variations were non-conclusive.

In this study, we established a foundation for further model-based analysis for uncertainty propagation that can be generalized to any kind of experiments in biogeosciences. Extensions comprising time-varying uncertainties by introducing a new random value for parameters at every time step or including covariance matrices, showing the simultaneous interaction of variations in two factors, can be straightforward implemented (de Castro, 2017). Finally, we believe that an explicit description of uncertainty quantification is essential for the interpretation and generalization of experimental results.

*Data availability.* Experimental data are available via the data portal Pangaia (PeECE II team, 2003; PeECE III team, 2005; Paul et al., 2014).

## Appendix A: Definition of relative growth rate

Relative growth rate  $\mu$  is calculated from the primary production rate by subtracting respiration and mortality losses as follows:  $\mu = P - R - L$ .

### A1 Primary production

Primary production rate reflects the limiting effects of light, dissolved inorganic carbon (DIC), temperature, and nutrient internal quota as follows:

$$P = P_{\max} \cdot f_{\text{PAR}} \cdot f_{\text{CO}_2} \cdot f_T \cdot f_Q \cdot f_p. \quad (\text{A1})$$

$P_{\max}$  is the maximum primary production rate, (Table 2). Specific light limitation  $f_{\text{PAR}}$  depends on light and  $\text{CO}_2$ . For the attenuation coefficient  $a_z$ , we consider that in coastal regions light intensity is typically reduced to 1 % of its surface value at 5 m (Denman and Gargett, 1983) and we obtained  $a_z = 0.75 \text{ m}^{-1}$ . Next, PAR experienced by cells at mixed layer depth (MLD = 4.5 m, Engel et al., 2008), was calculated from the level of radiation at the water surface,  $\text{PAR}_0$  (see Appendix B), following an exponential decay described by the Lambert–Beer law

$$\text{PAR} = \text{PAR}_0 \int_0^{\text{MLD}} e^{-a_z \cdot z} dz. \quad (\text{A2})$$

The relationship between photosynthesis and irradiance can be formulated by referring to a cumulative one-hit Poisson distribution (Ley and Mauzerall, 1982; Dubinsky et al., 1986). With the temperature and carbon acquisition dependence, it yields

$$f_{\text{PAR}} = \left( 1 - e^{-\frac{a_{\text{PAR}} \cdot \text{PAR}}{P_{\max} \cdot f_{\text{CO}_2} \cdot f_T}} \right), \quad (\text{A3})$$

where  $a_{\text{PAR}}$  is the effective absorption related to the chloroplast cross section and saturation response time for receptors (Geider et al., 1998a; Wirtz and Pahlow, 2010); the carbon acquisition term  $f_{\text{CO}_2}$  is described in Sect. 2.1, Eq. ().

$f_T$  is the temperature dependence. We considered that all metabolic rates depend on protein folding that increases with rising temperature following the Arrhenius equation (Scalley and Baker, 1997) as described in Geider et al. (1998b) or Schartau et al. (2007)

$$f_T = e^{-E_a \cdot \left( \frac{1}{T} - \frac{1}{T_{\text{ref}}} \right)}, \quad (\text{A4})$$

with activation energy  $E_a = \frac{T_{\text{ref}}^2}{10} \cdot \log(Q_{10})$  as in Wirtz (2013), where we used  $Q_{10} = 1.88$  for phytoplankton (Eppley, 1972; Brush et al., 2002) and  $T_{\text{ref}}$  was the mean measured temperature (see Appendix B).

The allometric factor  $\alpha_Q$  quantifies the scaling relation of subsistence quota and cell size. We used the Droop dependency on nutrient N : C ratio (Droop, 1973), which has been recently mechanistically derived (Wirtz and Pahlow, 2010; Pahlow and Oschlies, 2013)

$$f_Q = \left( 1 - \frac{Q_{\text{subs}}}{Q} \right), \quad (\text{A5})$$

where  $Q = \frac{\text{Phy}_N}{\text{Phy}_C}$ . Its lower reference, the subsistence quota  $Q_{\text{subs}} = Q_{\text{subs}}^* \cdot e^{-\alpha_Q \cdot \ell}$ , is considered size-dependent and reflects a lower protein demand for uptake mechanisms in large cells (Litchman et al., 2007).

The last term in Eq. (A1) accounts for an energy allocation trade-off in phytoplankton cells: protein allocation for photosynthetic compounds such as RuBisCo and pigments,  $f_p$ , versus allocation for nutrient uptake,  $f_v$ , expressed by  $f_p + f_v = 1$  (Wirtz and Pahlow, 2010; Pahlow and Oschlies, 2013). We simplified the detailed partition models by setting the trait fractions as constant.

### A2 Respiratory cost and nutrient uptake rates

Efforts related to nutrient uptake  $V$  are represented by a respiration term. Other expenses such as biosynthetic costs are neglected (Pahlow, 2005). The respiration rate is then calculated as  $R = \zeta \cdot V$ , where  $\zeta$  expresses the specific respiratory cost of nitrogen assimilation (Raven, 1980; Aksnes and Egge, 1991; Pahlow, 2005). For simplicity, our model merges the set of potentially limiting nutrients (e.g., P, Si and N) to a single resource only, i.e., DIN. We follow Aksnes and Egge (1991) as described in Pahlow (2005) for the maximum uptake rate

$$V_{\max} = \frac{1}{\frac{1}{V_{\max}^* \cdot f_T} + \frac{1}{\text{Aff} \cdot \text{DIN}}}, \quad (\text{A6})$$

comprising the maximum uptake coefficient  $V_{\max}^*$  and nutrient affinity Aff. In addition to a temperature dependence of nutrient uptake as reported by Schartau et al. (2007), we assumed that respiratory costs decrease with increasing cell size (Edwards et al., 2012), which leads to an allometric scaling in nutrient uptake (Wirtz, 2013) with exponent  $\alpha_V$ . We also accounted for the static proteins allocation trade-offs between photosynthetic machinery,  $f_p$ , and nutrients uptake,  $f_v = 1 - f_p$ . Thus, the nutrient uptake term yields

$$V = (1 - f_p) \cdot V_{\max} \cdot e^{-\alpha_V \cdot \ell}. \quad (\text{A7})$$

### A3 Loss rates

To describe the loss rate of phytoplankton biomass, we used a density-dependent term

$$L = L^* \cdot (\text{Phy}_C + \text{DH}_C). \quad (\text{A8})$$

The resulting matter flux increases the biomass of detritus and heterotrophs (DH), and a fraction of it becomes a part of



the remineralizable pool. A temperature-dependent remineralization term (Schartau et al., 2007)

$$r = r^* \cdot f_T \quad (\text{A9})$$

describes any kind of DIN production, such as hydrolysis and remineralization of organic matter, excretion of ammonia directly by zooplankton, and rapid remineralization of fecal pellets produced also by the zooplankton. The other fraction of the non-phytoplanktonic biomass is removed by settling with a rate related to the sinking coefficient,  $s$ , shown in Tables 1 and 2. Our model was calibrated with experimental data from enclosed mesocosms where aquarium pumps ensured mixing. Therefore, we assumed that sufficiently wealthy organisms could achieve neutral buoyancy (Boyd and Gradmann, 2002), and thus sinking might not have directly affected the phytoplankton biomass.

## Appendix B: Forcings

We used measured aquatic  $\text{CO}_2$  and temperature per mesocosm and ambient PAR, as model inputs (see Fig. 9). For the two PeECE experiments, the photon flux density was measured by the Geophysical Institute of the University of Bergen. To calculate the surface radiation inside the mesocosms,  $\text{PAR}_0$ , we followed (Schulz et al., 2008) and considered that 80 % of incident PAR passed through the gas tight tents, of which up to 15 % penetrated to approximately 2.5 m depth, the center of the mixed surface layer in PeECE III. The daily carbon dioxide data were interpolated and PAR signal was filtered by singular spectrum analysis to avoid sudden changes that could be detrimental to the performance of the numerical calculation, since the Heun method requires differentiable functions.

## Appendix C: Definition of POC

The applied model equations attribute phytoplankton, detritus, and herbivorous heterotrophs to particulate organic matter. Measurements of particulate organic carbon also include some fractions of large bacterioplankton, carnivorous zooplankton, as well as extracellular gel particles such as transparent exopolymer particles. These additional organic contributions to POC measurements are not explicitly resolved in our model. Therefore, for comparisons between simulation results and observations, we redefine the raw data from PANGAEA, named  $\text{POC}'$  hereafter (dots in Figs. 2, 3, and 5 represent the already modified POC data). We used data of transparent exopolymer particles, TEP, from Egge et al. (2009) for PeECE III, such as here  $\text{POC} = \text{POC}' - \text{TEP}$ . For PeECE II, TEP data were not available. We used  $\text{POC} = \text{POC}' - \text{POC}''$ , where  $\text{POC}''$  is the difference between particle abundance, PA, of the Coulter counter measurements and the flow cytometry data in Engel et al. (2008):

$$\text{POC}'' = \beta \cdot (\text{PA}_{\text{Coulter counter}} - \text{PA}_{\text{flow cytometry}}). \quad (\text{C1})$$

The scaling parameter  $\beta = 0.000065 \mu\text{mol-CL}^{-1}$  was tuned to provide reductions between 40 and 50 % from total POC, in agreement with the adjustments of PeECE III.

## Appendix D: Model representation of replicates

Heuristic exploration of the potential origins of the observed variability uses statistical inference tools, such as a multi-way repeated measures ANOVA, exploring which independent factors are contributing the most to the standard deviations. Such approaches have the advantage of accounting for interacting effects between combinations of factors (and not only for the synergistic effects of each factor and acidification, as in our model-based approach; see Sect. 3). However, the realization through an experimental setup would make a high-dimensional multi-factorial experiment extremely difficult to perform (Fig. 8). For three acidification levels, the minimum number of factor levels (i.e., high and low), minimum number of sample units (i.e., duplicates), and the same number of factors we analyze here, (i.e.,  $N = 19$ ), the total number of mesocosms would be  $3 \times 2 \times 2 \times 19 = 228$ . The possibility of simulating a high number of replicates is one of the unique strengths of modeling. For each factor, we simulate possible realizations of the same acidification level with slight variations of the factor reference value (simulating differences in physiological states and community structure). We generated model solutions for  $10^4$  normally distributed factor values, i.e., in total, 3 acidification levels  $\times$  19 factors  $\times$   $10^4$  virtual replicates for PeECE II and III experiments. Examples of 50 virtual replicates with uncertainty in initial nutrient concentration are shown in Fig. 8 and examples of 10 virtual replicates with uncertainty in phytoplankton biomass losses are shown in Fig. 1, both numerically calculated for low  $\text{CO}_2$  conditions in PeECE III.

## Appendix E: Residuals of the model–data fit

For the model–data fit shown in Figs. 2 and 3, we calculated the cumulative residuals  $E$  and  $M$  (Table E1) with respect to the mean of experimental replicates per treatment, time, and mesocosm. For experimental data, residuals  $E$  were calculated as follows:

$$E = \sum_{\text{treat,rep,day}} |Y_{\text{treat,rep,day}}^{\text{exp}} - \langle Y_{\text{treat,day}}^{\text{exp}} \rangle| / \eta \quad (\text{E1})$$

and for model results, residuals  $M$  were calculated as follows:

$$M = \sum_{\text{treat,rep,day}} |Y_{\text{treat,rep,day}}^{\text{mod}} - \langle Y_{\text{treat,day}}^{\text{exp}} \rangle| / \eta \quad (\text{E2})$$

with  $\eta = 9$  being the total number of mesocosms. High residuals entail high deviation from the trend. In the case of  $E$ , this is the deviation from the mean of the treatment (typically used in statistical inference tools), and in the case of  $M$ , the deviation from the model reference run. When both  $E$  and  $M$  values are comparable, we can infer that the quality of both representations is similar (see Table E1). Thus, conclusions inferred from both approaches are based on equally valid assumptions.

**Table E1.** Cumulative residuals for PeECE III.

$Y$	$E$	$M$	units
POC	35.1	37.4	$\mu\text{mol-CL}^{-1}$
PON	6.0	9.1	$\mu\text{mol-NL}^{-1}$
DIN	6.7	9.2	$\mu\text{mol-NL}^{-1}$

**Author contributions.** Kai Wirtz, Markus Schartau, and Maria Moreno de Castro developed the model code; Maria Moreno de Castro performed the simulations and prepared the manuscript, which was revised by Kai Wirtz and Markus Schartau.

**Competing interests.** The authors declare that they have no conflict of interest.

**Acknowledgements.** We thank Sabine Mathesius for the PAR and temperature data for both the PeECE II and III experiments and Kaela Slavik for the English edition of the preliminary version of the manuscript. We acknowledge our two anonymous reviewers for their helpful comments and suggestions. This work is a contribution to the National German project Biological Impacts of Ocean Acidification (BIOACID) and it is also supported by the Helmholtz society via the program PACES.

The article processing charges for this open-access publication were covered by a Research Centre of the Helmholtz Association.

Edited by: M. Grégoire

Reviewed by: two anonymous referees

## References

- Adamson, M. and Morozov, A.: Defining and detecting structural sensitivity in biological models: developing a new framework, *J. Math. Biol.*, 69, 1815–1848, doi:10.1007/s00285-014-0753-3, 2014.
- Aksnes, D. L. and Egge, J. K.: A theoretical model for nutrient uptake in phytoplankton, *Mar. Ecol. Prog. Ser.*, 70, 65–72, 1991.
- Antia, N. J., MacAllister, C. D., Parsons, T. R., Stephens, K., and Strickland, J. D. H.: Further measurements of primary production using a large-volume plastic sphere, *Limnol. Oceanogr.*, 8, 166–173, doi:10.4319/lo.1963.8.2.0166, 1963.
- Artioli, Y., Blackford, J. C., Nondal, G., Bellerby, R. G. J., Wake-  
lin, S. L., Holt, J. T., Butenschön, M., and Allen, J. I.: Heterogeneity of impacts of high CO<sub>2</sub> on the North Western European Shelf, *Biogeosciences*, 11, 601–612, doi:10.5194/bg-11-601-2014, 2014.
- Biddanda, B. and Benner, R.: Carbon, nitrogen, and carbohydrate fluxes during the production of particulate and dissolved organic matter by marine phytoplankton, *Limnol. Oceanogr.*, 42, 506–518, doi:10.4319/lo.1997.42.3.0506, 1997.
- Paul, C., Matthiessen, B., and Sommer, U.: Mesocosm experiment 2012 on warming and acidification effects on phytoplankton biomass and chemical composition, PANGAEA, available at: doi:10.1594/PANGAEA.840852, 2014.
- Boyd, C. M. and Gradmann, D.: Impact of osmolytes on buoyancy of marine phytoplankton, *Mar. Biol.*, 141, 605–618, 2002.
- Brennan, A.: Necessary and Sufficient Conditions, in: *The Stanford Encyclopedia of Philosophy*, edited by: Zalta, E. N., spring 2012 edn., 2012.
- Broadgate, W., Riebesell, U., Armstrong, C., Brewer, P., Denman, K., Feely, R., Gao, K., Gatusso, J. P., Isensee, K., Kleypas, J., Laffoley, D., Orr, J., Pöetner, H. O., de Rezende, C. E., Schimdt, D., Urban, E., Waite, A., and Valdés, L.: Ocean acidification summary for policymakers – Third Symposium on the ocean in a high-CO<sub>2</sub> world, International Geosphere-Biosphere Programme, Sweden, p. 26, 2013.
- Brush, M., Brawley, J., Nixon, S., and Kremer, J.: Modeling phytoplankton production: problems with the Eppley curve and an empirical alternative, *Mar. Ecol. Prog. Ser.*, 238, 31–45, doi:10.3354/meps238031, 2002.
- Caldeira, K. and Wickett, M. E.: Oceanography: Anthropogenic carbon and ocean pH, *Nature*, 425, 365–365, doi:10.1038/425365a, 2003.
- Chantrasmí, T. and Iaccarino, G.: Forward and backward uncertainty propagation for discontinuous system response using the Pade-Legendre method, *International Journal for Uncertainty Quantification*, 2, 125–143, 2012.
- Chen, C. Y.: Effect of pH on the growth and carbon uptake of marine phytoplankton, *Mar. Ecol. Prog. Ser.*, 109, 83–94, 1994.
- Cohen, J.: *Statistical Power Analysis for the Behavioral Sciences*, Lawrence Erlbaum Associates, Hillsdale, NJ, 2nd edn., 1988.
- Cornwall, C. and Hurd, C.: Experimental design in ocean acidification research: problems and solutions, *ICES Journal of Marine Science*, 73, 572–581, doi:10.1093/icesjms/fsv118, 2015.
- Cottingham, K. L., Lennon, J. T., and Brown, B. L.: Knowing when to draw the line: designing more informative ecological experiments, *Front. Ecol. Environ.*, doi:10.1890/1540-9295(2005)003[0145:KWTDTL]2.0.CO;2, 2005.
- Denman, K. L. and Gargett, A. E.: Time and space scales of vertical mixing and advection of phytoplankton in the upper ocean, *Limnol. Oceanogr.*, 28, 801–815, 1983.
- Droop, M. R.: Some thoughts on nutrient limitation in algae, *J. Phycol.*, 9, 264–272, doi:10.1111/j.1529-8817.1973.tb04092.x, 1973.
- Dubinsky, Z., Falkowski, P. G., and Wyman, K.: Light harvesting and utilization by phytoplankton, *Plant Cell Physiol.*, 21, 1335–1349, 1986.
- Edwards, K., Klausmeier, C. A., and Litchman, E.: Allometric scaling and taxonomic variation in nutrient utilization traits and maximum growth rate of phytoplankton, *Limnol. Oceanogr.*, 57, 554–556, 2012.
- Egge, J. K., Thingstad, T. F., Larsen, A., Engel, A., Wohlers, J., Bellerby, R. G. J., and Riebesell, U.: Primary production during nutrient-induced blooms at elevated CO<sub>2</sub> concentrations, *Biogeosciences*, 6, 877–885, doi:10.5194/bg-6-877-2009, 2009.
- Eggers, S. L., Lewandowska, A. M., Barcelos e Ramos, J., Blanco-Ameijeiras, S., Gallo, F., and Matthiessen, B.: Community composition has greater impact on the functioning of marine phytoplankton communities than ocean acidification, *Glob. Change Biol.*, 20, 713–723, doi:10.1111/gcb.12421, 2014.
- Ellison, S. L. R. and Williams, A.: *Eurachem/CITAC guide: Quantifying Uncertainty in Analytical Measurement*, third edn., p. 26, 2012.
- Engel, A., Schulz, K. G., Riebesell, U., Bellerby, R., Delille, B., and Schartau, M.: Effects of CO<sub>2</sub> on particle size distribution and phytoplankton abundance during a mesocosm bloom experiment (PeECE II), *Biogeosciences*, 5, 509–521, doi:10.5194/bg-5-509-2008, 2008.
- Engel, A., Cisternas Novoa, C., Wurst, M., Endres, S., Tang, T., Schartau, M., and Lee, C.: No detectable effect of CO<sub>2</sub> on el-

- emental stoichiometry of *Emiliana huxleyi* in nutrient-limited, acclimated continuous cultures, *Mar. Ecol. Prog. Ser.*, 507, 15–30, 2014.
- Engel, A., Zondervan, I., Aerts, K., Beaufort, L., Benthien, A., Chou, L., Belille, B., Gattuso, J.-P., Harlay, J., Heemann, C., Hoffmann, L., Jacquet, S., Nejstgaard, J., Pizay, M.-D., Rochelle-Newall, E., Scheider, U., Terbruggen, A., and Riebesell, U.: Testing the direct effect of CO<sub>2</sub> concentration on a bloom of the coccolithophorid *Emiliana huxleyi* in mesocosm experiments, *Limnol. Oceanogr.*, 50, 493–507, 2005.
- Eppley, R. W.: Temperature and phytoplankton growth in the sea, *Fishery Bulletin*, 1972.
- Field, A., Miles, J., and Field, Z.: *Discovering statistics using R*, SAGE Publications Ltd, 2008.
- Fussmann, G. F. and Blasius, B.: Community response to enrichment is highly sensitive to model structure, *Biol. Lett.*, 1, 9–12, doi:10.1098/rsbl.2004.0246, 2005.
- Gao, K., Helbling, E. W., Häder, D. P., and Hutchins, D. A.: Responses of marine primary producers to interactions between ocean acidification, solar radiation, and warming, *Mar. Ecol. Prog. Ser.*, 470, 167–189, doi:10.3354/meps10043, 2012.
- Geider, R., Macintyre, Graziano, L., and McKay, R. M.: Responses of the photosynthetic apparatus of *Dunaliella tertiolecta* (Chlorophyceae) to nitrogen and phosphorus limitation, *European Journal of Phycology*, 33, 315–332, doi:10.1080/09670269810001736813, 1998a.
- Geider, R. J., MacIntyre, H. L., and Kana, T. M.: A dynamic regulatory model of phytoplankton acclimation to light, nutrients, and temperature, *Limnol. Oceanogr.*, 43, 679–694, doi:10.4319/lo.1998.43.4.0679, 1998b.
- JCGM: Guide to the Expression of Uncertainty in Measurement (GUM 1995 with minor corrections) by a Joint Committee for Guides in Metrology (JCGM 100:2008), available at: [http://www.bipm.org/utis/common/documents/jcgm/JCGM\\_100\\_2008\\_E.pdf](http://www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf), 2008a.
- JCGM: Supplement 1 to the 'Guide to the Expression of Uncertainty in Measurement – Propagation of distributions using a Monte Carlo method (JCGM 101:2008), available at: [http://www.bipm.org/utis/common/documents/jcgm/JCGM\\_101\\_2008\\_E.pdf](http://www.bipm.org/utis/common/documents/jcgm/JCGM_101_2008_E.pdf), 2008b.
- Jones, B. M., Iglesias-Rodriguez, M. D., Skipp, P. J., Edwards, R. J., Greaves, M. J., Jeremy, R. Y., Elderfield, H., and O'Connor, D.: Responses of the *Emiliana huxleyi* Proteome to Ocean Acidification, *PLoS ONE*, 8, 2857–2869, doi:10.1371/journal.pone.0061868, 2014.
- Kennedy, M. C. and O'Hagan, A.: Bayesian Calibration of Computer Models, *Journal of the Royal Statistical Society, Series B*, 63, 425–464, 2001.
- Kim, J.-M., Lee, K., Shin, K., Kang, J.-H., Lee, H.-W., Kim, M., Jang, P.-G., and Jang, M.-C.: The effect of seawater CO<sub>2</sub> concentration on growth of a natural phytoplankton assemblage in a controlled mesocosm experiment, *Limnol. Oceanogr.*, 51, 1629–1636, 2006.
- Kroeker, K. J., Kordas, R. L., Crim, R., Hendriks, I. E., Ramajo, L., Singh, G. S., Duarte, C. M., and Gattuso, J.-P.: Impacts of ocean acidification on marine organisms: quantifying sensitivities and interaction with warming, *Glob. Change Biol.*, 19, 1884–1896, doi:10.1111/gcb.12179, 2013.
- Larssen, T., Huseby, R. B., Cosby, B. J., Høst, G., Høgåsen, T., and Aldrin, M.: Forecasting acidification effects using a Bayesian calibration and uncertainty propagation approach, *Environ. Sci. Technol.*, 40, 7841–7847, 2006.
- Ley, A. C. and Mauzerall, D. C.: Absolute absorption cross-sections for photosystem II and the minimum quantum requirement for photosynthesis in *Chlorella vulgaris*, *Biochimica et Biophysica Acta*, 680, 95–106, 1982.
- Litchman, E., Klausmeier, C. A., Schofield, O., and Falkowski, P.: The role of functional traits and trade-offs in structuring phytoplankton communities: scaling from cellular to ecosystem level, *Ecol. Lett.*, 10, 1170–1181, 2007.
- Miller, R. G. J.: *Beyond ANOVA, Basics of Applied Statistics*, Wiley, New York – Chichester – Brisbane – Toronto – Singapore, 1988.
- Moreno de Castro, M.: Tolerance of mesocosm experiments to time-varying uncertainties, in preparation, 2017.
- Nagelkerken, I. and Connell, S. D.: Global alteration of ocean ecosystem functioning due to increasing human CO<sub>2</sub> emissions, *P. Natl. Acad. Sci.*, 112, 13272–13277, doi:10.1073/pnas.1510856112, 2015.
- Pahlow, M.: Linking chlorophyll–nutrient dynamics to the Redfield N:C ratio with a model of optimal phytoplankton growth, *Mar. Ecol. Prog. Ser.*, 287, 33–43, 2005.
- Pahlow, M. and Oschlies, A.: Optimal allocation backs Droop's cell-quota model, *Mar. Ecol. Prog. Ser.*, 473, 1–5, 2013.
- PeECE II team: PeECE II – Pelagic Ecosystem CO<sub>2</sub> Enrichment Study, Raunefjord, Bergen, Norway, 2003, PANGAEA, available at: doi:10.1594/PANGAEA.723045, 2003.
- PeECE III team: PeECE II – Pelagic Ecosystem CO<sub>2</sub> Enrichment Study, Raunefjord, Bergen, Norway, 2005, PANGAEA, available at: doi:10.1594/PANGAEA.726955, 2005.
- Paul, C., Matthiessen, B., and Sommer, U.: Warming, but not enhanced CO<sub>2</sub> concentration, quantitatively and qualitatively affects phytoplankton biomass, *Mar. Ecol. Prog. Ser.*, 528, 39–51, doi:10.3354/meps11264, 2015.
- Peterman, R. M.: The importance of reporting statistical power: the forest decline and acidic deposition example, *Ecology*, 71, 2024–2027, 1990.
- R Core Team: *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, available at: <https://www.R-project.org/> (last access: 3 April 2017), 2016.
- Raven, J. and Beardall, J.: Carbon Acquisition Mechanisms of Algae: Carbon Dioxide Diffusion and Carbon Dioxide Concentrating Mechanisms, in: *Photosynthesis in Algae*, edited by: Larkum, A., Douglas, S., and Raven, J., vol. 14 of *Advances in Photosynthesis and Respiration*, 225–244, Springer Netherlands, doi:10.1007/978-94-007-1038-2\_11, 2003.
- Raven, J. A.: Nutrient transport in microalgae, *Adv. Microb. Physiol.*, 21, 47–226, 1980.
- Riebesell, U. and Tortell, P. D.: Effects of Ocean Acidification on Pelagic Organisms and Ecosystems, in: *Ocean Acidification*, edited by: Gattuso, J.-P. and Hansson, L., 99–121, Oxford University Press, Oxford, UK, 2011.
- Riebesell, U., Wolf-Gladrow, D. A., and Smetacek, V.: Carbon dioxide limitation of marine phytoplankton growth rates, *Nature*, 361, 249–251, doi:10.1038/361249a0, 1993.

- Riebesell, U., Zondervan, I., Rost, B., Tortell, P. D., Zeebe, R. E., and Morel, F. M. M.: Reduced calcification of marine plankton in response to increased atmospheric  $\text{CO}_2$ , *Nature*, 407, 364–367, doi:10.1038/35030078, 2000.
- Riebesell, U., Schulz, K. G., Bellerby, R. G. J., Botros, M., Fritzsche, P., Meyerhöfer, M., Neill, C., Nondal, G., Oschlies, A., Wohlers, J., and Zöllner, E.: Enhanced biological carbon consumption in a high  $\text{CO}_2$  ocean, *Nature*, 450, 545–548, doi:10.1038/nature06267, 2007.
- Riebesell, U., Bellerby, R. G. J., Grossart, H.-P., and Thingstad, F.: Mesocosm  $\text{CO}_2$  perturbation studies: from organism to community level, *Biogeosciences*, 5, 1157–1164, doi:10.5194/bg-5-1157-2008, 2008.
- Riebesell, U., Fabry, V. J., Hansson, L., and Gattuso, J. P.: Guide to best practices for ocean acidification research and data reporting, Publications Office of the European Union, 2010.
- Rost, B., Riebesell, U., Burkhardt, S., and Suetemeyer, D.: Carbon acquisition of bloom-forming marine phytoplankton, *Limnol. Oceanogr.*, 48, 55–67, 2003.
- Ruxton, G. D. and Colegrave, N.: Experimental design for the life sciences, Oxford: Oxford University Press, 2006.
- Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., Wanninkhof, R., Wong, C. S., Wallace, D. W. R., Tilbrook, B., Millero, F. J., Peng, T.-H., Kozyr, A., Ono, T., and Rios, A. F.: The Oceanic Sink for Anthropogenic  $\text{CO}_2$ , *Science*, 305, 367–371, doi:10.1126/science.1097403, 2004.
- Scalley, M. L. and Baker, D.: Protein folding kinetics exhibit an Arrhenius temperature dependence when corrected for the temperature dependence of protein stability, *P. Natl. Acad. Sci.*, 94, 10636–10640, doi:10.1073/pnas.94.20.10636, 1997.
- Schartau, M., Engel, A., Schröter, J., Thoms, S., Völker, C., and Wolf-Gladrow, D.: Modelling carbon overconsumption and the formation of extracellular particulate organic carbon, *Biogeosciences*, 4, 433–454, doi:10.5194/bg-4-433-2007, 2007.
- Scheinin, M., Riebesell, U., Rynearson, T. A., Lohnbeck, K. T., and Collins, S.: Experimental evolution gone wild, *J. R. Soc. Interface*, 12, doi:10.1098/rsif.2015.0056, 2015.
- Schluter, L., Lohnbeck, K. T., Gutowska, M. A., Groger, J. A., Riebesell, U., and Reusch, T. B. H.: Adaptation of a globally important coccolithophore to ocean warming and acidification, *Nature Climate Change*, 4, 1024–1030, doi:10.1038/nclimate2379, 2014.
- Schulz, K. G., Riebesell, U., Bellerby, R. G. J., Biswas, H., Meyerhöfer, M., Müller, M. N., Egge, J. K., Nejstgaard, J. C., Neill, C., Wohlers, J., and Zöllner, E.: Build-up and decline of organic matter during PeECE III, *Biogeosciences*, 5, 707–718, doi:10.5194/bg-5-707-2008, 2008.
- Sommer, U., Paul, C., and Moustaka-Gouni, M.: Warming and Ocean Acidification Effects on Phytoplankton – From Species Shifts to Size Shifts within Species in a Mesocosm Experiment, *PLOS ONE*, 10, 39–51, doi:10.1371/journal.pone.0125239, 2015.
- Tanaka, T., Thingstad, T. F., Løvdal, T., Grossart, H.-P., Larsen, A., Allgaier, M., Meyerhöfer, M., Schulz, K. G., Wohlers, J., Zöllner, E., and Riebesell, U.: Availability of phosphate for phytoplankton and bacteria and of glucose for bacteria at different  $p\text{CO}_2$  levels in a mesocosm study, *Biogeosciences*, 5, 669–678, doi:10.5194/bg-5-669-2008, 2008.
- Toral, R. and Colet, P.: Stochastic Numerical Methods, Wiley-VCH, 2014.
- Tortell, P. D., Payne, C. D., Li, Y., Trimborn, S., Rost, B., Smith, W. O., Riesselman, C., Dunbar, R. B., Sedwick, P., and DiTullio, G. R.:  $\text{CO}_2$  sensitivity of Southern Ocean phytoplankton, *Geophys. Res. Lett.*, 35, 104605, doi:10.1029/2007GL032583, 2008.
- Wirtz, K. W.: Non-uniform scaling in phytoplankton growth rate due to intracellular light and  $\text{CO}_2$  decline, *J. Plankton Res.*, 33, 1325–1341, 2011.
- Wirtz, K. W.: Mechanistic origins of variability in phytoplankton dynamics: Part I: Niche formation revealed by a Size-Based Model, *Mar. Biol.*, 160, 2319–2335, 2013.
- Wirtz, K. W. and Pahlow, M.: Dynamic chlorophyll and nitrogen:carbon regulation in algae optimizes instantaneous growth rate, *Mar. Ecol. Prog. Ser.*, 402, 81–96, 2010.
- Zondervan, I., Zeebe, R. E., Rost, B., and Riebesell, U.: Decreasing marine biogenic calcification: A negative feedback on rising atmospheric  $p\text{CO}_2$ , *Global Biogeochem. Cy.*, 15, 507–516, doi:10.1029/2000GB001321, 2001.